

# Phase retrieval from single biomolecule diffraction pattern

Shiro Ikeda\*

Hidetoshi Kono†

December 12, 2011

## Abstract

In this paper, we propose the SPR (sparse phase retrieval) method, which is a new phase retrieval method for coherent x-ray diffraction imaging (CXDI). Conventional phase retrieval methods effectively solve the problem for high signal-to-noise ratio measurements, but would not be sufficient for single biomolecular imaging which is expected to be realized with femto-second x-ray free electron laser pulses. The SPR method is based on the Bayesian statistics. It does not need to set the object boundary constraint that is required by the commonly used hybrid input-output (HIO) method, instead a prior distribution is defined with an exponential distribution and used for the estimation. Simulation results demonstrate that the proposed method reconstructs the electron density under a noisy condition even some central pixels are masked.

## 1 Introduction

X-ray free electron lasers (XFELs) can potentially provide us a novel mean to determine the three-dimensional (3D) structure of biomolecules from the diffraction images of single molecules [1, 2]. Conventionally, x-ray crystallography has been the principal tool to determine high-resolution 3D structures of proteins, nucleic acids, and their complexes. However, a critical process is the crystallization, where a sufficiently large crystal must be prepared. It is known that about 40% of biomolecules, particularly membrane proteins, do not crystallize. The coherent x-ray diffraction imaging (CXDI) by XFELs does not require a crystal and will change the situation.

In order to realize the single molecule imaging, a straightforward scenario has been proposed [3, 4] and demonstrated with single mimivirus particles [5]. Firstly, a series of single molecules with the same structure are injected into vacuum and exposed to a very short (less than 5 fs) pulse x-ray beam. The diffraction image of the molecule with unknown orientation is recorded before radiation-induced structural degradation. This process is repeated until a sufficient number of images are recorded. The diffraction data are then classified into groups according to the orientations of the molecule and averaged to improve the signal-to-noise ratio [6, 7]. Next, the relative orientations of the classes are determined to assemble a 3D coherent diffraction pattern in reciprocal space. Finally, the 3D electron density of the molecule is obtained by a phase retrieval method.

The above scenario was proved to be effective for the 3D reconstruction of mimivirus particles [5]. However, the condition for a single molecule CXDI will be more severe since the number of the photons received by each pixel will be very limited. Some methods of reconstructing 3D diffraction intensity have been proposed for this problem [8–10]. They are likely to be computationally intensive and the feasibility remains to be seen.

Another scenario might be possible. After collecting a sufficient number of diffraction images with unknown orientations, 2D electron density images are first reconstructed from the 2D diffraction data, followed by the 3D reconstruction method used in transmission electron microscopy [11]. In either scenario, the difficulty is due to the small number of photons scattered by a single molecule even though

---

\*The Institute of Statistical Mathematics, Tokyo, Japan, shiro@ism.ac.jp

†Japan Atomic Energy Agency, Kyoto, Japan, kono.hidetoshi@jaea.go.jp

x-ray beams that linac coherent light source (LCLS) has started to commission give us illumination which is around  $10^9$  times brighter ( $\sim 10^{19}$  photons/pulse/mm<sup>2</sup> in the unfocused beam) than that of current synchrotrons [12].

We propose a new method for retrieving phases of a diffraction image which will be obtained by XFELs. The method is based on the maximum a posteriori (MAP) estimation of the Bayesian statistics. The posterior distribution is computed from the likelihood and the prior. For the present problem, the Poisson distribution is appropriate for the likelihood. The prior reflects the assumption of the electron density. Since the electron density is known to have a lot of zeros, we employ the exponential distribution which encourages the MAP estimate to have a lot of zeros (this is called a sparse solution [13, 14]. A similar idea is employed in [15]). Therefore, we call our method, sparse phase retrieval (SPR) method.

This combination is effective for the problem and is distinct from other Bayesian approaches proposed for different phase retrieval problems [16, 17]. Compared to the widely used hybrid input-output (HIO) method [18–20], the SPR method does not require the support region and is robust against the limited photon counts and the missing central pixels. In this paper, we focus on 2D phase retrieval problem. The resulting SPR method can be used directly for the second scenario. If 3D diffraction data are available, it is also possible to apply the SPR method. But reconstruction of 3D diffraction data from 2D images with unknown orientations is beyond the scope of the present paper.

## 2 Proposed method

### 2.1 Diffraction image and noise

We first formulate the phase retrieval problem of a single diffraction image. Let  $f(x, y) \geq 0$  be the electron density of a molecule projected onto a two-dimensional (2D) plane. The coordinate is discretized, that is,  $x, y = 1, \dots, M$ . We introduce the following notation

$$\mathbf{f} = \{f(x, y)\} = \{f_{xy}\} = \{f_{11}, \dots, f_{MM}\}, \quad \mathbf{F} = \{F(u, v)\} = \{F_{uv}\} = \{F_{11}, \dots, F_{MM}\},$$

where,  $\mathbf{F}$  is the Fourier transform of  $\mathbf{f}$  defined as follows,

$$F_{uv} = (\mathcal{F}(\mathbf{f}))_{uv} = \frac{1}{M} \sum_{x,y} f_{xy} \exp\left(\frac{2\pi i(ux + vy)}{M}\right). \quad (1)$$

The coefficient of the Fourier transform differs from the commonly used  $1/M^2$ . We prefer the above definition because the power of  $\mathbf{F}$  and  $\mathbf{f}$  is the same. We assume that the frame size is sufficiently larger than the molecule in the image and many components of  $\mathbf{f}$  are zero.

In the ideal situation, each pixel of a diffraction image receives the number of photons proportional to the power spectrum  $|F_{uv}|^2$ , where  $(u, v)$  is the 2D index of the pixel. However, the XFEL measurement of a single molecule diffraction will be different because the flux of the x-ray laser is not sufficiently large. The number of the photons detected by each sensor will be small and behave stochastically.

Let  $N_{uv}$  be the number of photons detected by the sensor at  $(u, v)$ . We study the distribution of  $N_{uv}$ . The total number  $N_{all} = \sum_{uv} N_{uv}$  is a stochastic variable which follows the Poisson distribution. The expected value of  $N_{all}$  is proportional to the intensity  $I_X$  of x-ray. Assuming each  $N_{uv}$  is independent and the distribution belongs to the same family,  $N_{uv}$  also follows a Poisson distribution. Let the expected value of  $N_{uv}$  be  $S_{uv}$  and the distribution of  $N_{uv}$  becomes

$$p(N_{uv}|S_{uv}) = \frac{S_{uv}^{N_{uv}} \exp(-S_{uv})}{N_{uv}!}.$$

We only consider this counting process and do not consider other observation noise. Approximately,  $S_{uv}$  is denoted as  $S_{uv} = \alpha |F_{uv}|^2 \cos^3 \theta$ , where  $\alpha$  is a positive constant and  $\theta$  is the scattering angle (see Fig. 1). Because  $\theta$  is a function of  $u$  and  $v$ , we further rewrite it as  $\alpha |F_{uv}|^2 c_{uv}$ , where  $c_{uv} = c(u, v) = \cos^3 \theta$ .

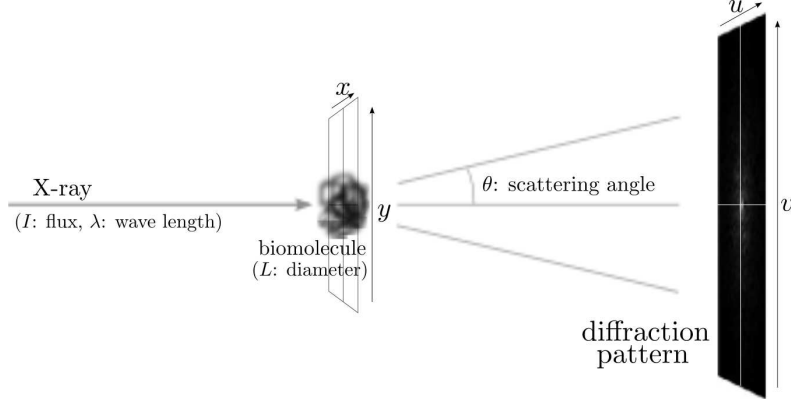


Figure 1: A schematic picture of the single biomolecular diffraction imaging with XFEL.

The Fourier transform,  $\mathbf{F}$ , is a linear function of  $\mathbf{f}$ , and the probability of  $\mathbf{N}$ , where  $\mathbf{N} = \{N_{uv}\} = \{N_{11}, \dots, N_{MM}\}$ , is denoted as follows

$$p(\mathbf{N}|\mathbf{f}) = \prod_{uv} \frac{(|F_{uv}|^2 c_{uv})^{N_{uv}} \exp(-|F_{uv}|^2 c_{uv})}{N_{uv}!}. \quad (2)$$

Note that the scaling of  $|F_{uv}|^2$  cannot be determined only from the observed scattering pattern and  $\alpha$  is set to 1. This scaling indeterminacy is ignored here. It would be recovered from other knowledge, such as the total energy.

In the following, we consider the density reconstruction as an estimation of  $\mathbf{f}$ . Commonly used estimate is the maximum likelihood estimate (MLE) which maximizes  $p(\mathbf{N}|\mathbf{f})$ . The likelihood in Eq.(2) is maximized by setting  $|F_{uv}|^2 = N_{uv}/c_{uv}$ , but this is not sufficient for the estimation of  $\mathbf{f}$  because the phase is lost. Some additional information is needed.

## 2.2 SPR method

Instead of the MLE, the MAP estimate of the Bayesian statistics is used for the SPR method. The MAP estimate maximizes the posterior distribution, where the posterior distribution is defined as the product of the likelihood function and the prior distribution. Similar framework has been proposed for different phase retrieval problems [16, 17].

The likelihood function is defined as a Poisson distribution as in Eq.(2) and we discuss the prior distribution in this subsection. For a 2D image of a single molecule diffraction, it is natural to assume that the electron density is high in the central part and becomes zero in the peripheral part. But we do not know which pixel is zero. The prior should reflect this knowledge.

It is known that the prior distribution with an exponential distribution encourages the MAP estimate to have a lot of zeros [13, 14] and we use this for the SPR method

$$p(\mathbf{f}) = \prod_{xy} p(f_{xy}), \quad p(f_{xy}) = \rho_{xy} \exp(-\rho_{xy} f_{xy}), \quad f_{xy} \geq 0, \quad \rho_{xy} \geq 0,$$

where each  $f_{xy}$  is assumed to be independent. The hyperparameter  $\rho_{xy}$  reflects the prior belief of  $f_{xy} = 0$ . Larger the value of  $\rho_{xy}$ , stronger the belief. In order to reflect the prior knowledge, we define the following one parameter function  $\rho(\mu)_{xy}$

$$\rho(\mu)_{xy} = \mu w_{xy}, \quad w_{xy} = a \left\{ \left( x - \frac{1+M}{2} \right)^2 + \left( y - \frac{1+M}{2} \right)^2 \right\} + b. \quad (3)$$

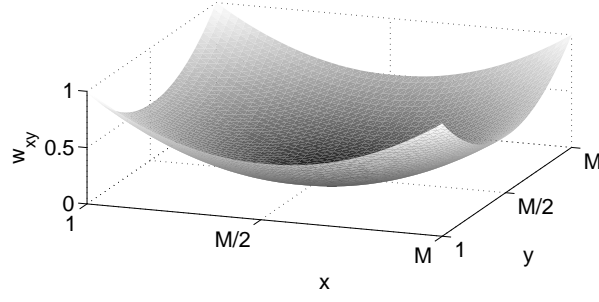


Figure 2: The 3D plot of  $w_{xy}$ . This reflects the prior knowledge of the electron density.

The two parameters  $a$  and  $b$  were adjusted to make  $w_{xy}$  equal to 0 at the center and 1 at the corners (see Fig. 2) and  $\mu$  adjusts the overall strength of the prior. We have employed a parabolic function, but if there is any prior knowledge,  $\rho(\mu)_{xy}$  should reflect it.

The posterior distribution of  $\mathbf{f}$  observing  $\mathbf{N}$  has the following relation,

$$p(\mathbf{f}|\mathbf{N}) \propto p(\mathbf{N}|\mathbf{f})p(\mathbf{f}).$$

It is convenient to take the logarithm of  $p(\mathbf{f}|\mathbf{N})$  in order to compute the MAP estimate. By taking the logarithm and collecting terms related to  $\mathbf{f}$ , the following function is obtained

$$\ell_{\mu}(\mathbf{f}|\mathbf{N}) = \sum_{uv} (N_{uv} \ln |F_{uv}|^2 - |F_{uv}|^2 c_{uv}) - \mu \sum_{xy} w_{xy} f_{xy}. \quad (4)$$

The first summation corresponds to the likelihood with the Poisson distribution (similar idea is used in [16]), the second summation represents the prior with the exponential distribution, and  $\mu$  controls the balance between two terms. This combination is the key idea of the proposed method. The SPR method is summarized as follows.

$$\text{maximize } \ell_{\mu}(\mathbf{f}|\mathbf{N}) \quad \text{subject to } f_{xy} \geq 0 \quad (1 \leq x, y \leq M). \quad (5)$$

We start with a large  $\mu$  and decrease it. The MAP estimate is computed for each  $\mu$ , and the best value of  $\mu$  is selected according to some criterion. This issue will be discussed later.

We note that our formulation is related to the compressed sensing (CS) [21], and the idea of CS has been applied to the phase retrieval problem [15]. Our approach is different from the work since the Poisson likelihood is included.

### 2.3 Algorithm

A gradient based algorithm is used to compute the MAP estimate. The algorithm is shown below starting with the following relation,

$$\frac{\partial |F_{uv}|^2}{\partial f_{xy}} = \frac{2}{M} \text{Re} \left( F_{uv} \exp \left( -\frac{2\pi i (ux + vy)}{M} \right) \right).$$

By defining the inverse Fourier transform as

$$(\mathcal{F}^{-1}(\mathbf{F}))_{xy} = \frac{1}{M} \sum_{uv} F_{uv} \exp \left( -\frac{2\pi i (ux + vy)}{M} \right),$$

the derivative of  $\ell_{\mu}(\mathbf{f}|\mathbf{N})$  becomes

$$\frac{\partial \ell_{\mu}(\mathbf{f}|\mathbf{N})}{\partial f_{xy}} = 2 \text{Re}(\mathcal{F}^{-1}(\mathbf{g}(\mathbf{F}; \mathbf{N})))_{xy} - \mu w_{xy},$$

where  $\mathbf{g}(\mathbf{F}; \mathbf{N}) = \{g_{uv}(F_{uv}; N_{uv})\}$  and

$$g_{uv}(F_{uv}; N_{uv}) = \left( \frac{N_{uv}}{|F_{uv}|^2} - c_{uv} \right) F_{uv}.$$

In order to compute the derivative of  $\ell_\mu(\mathbf{f}|\mathbf{N})$ ,  $\mathcal{F}$  is applied to  $\mathbf{f}$ , then  $\mathcal{F}^{-1}$  is applied to  $\mathbf{g}(\mathbf{F}; \mathbf{N})$ . Thus, the computational cost is similar to the HIO method [19].

We employed a naïve iterative updating rule for the MAP estimate of  $f_{xy}$ ,  $x, y = 1, \dots, M$ .

$$f_{xy}^{t+1} = \max\left(0, f_{xy}^t + \eta_t \frac{\partial \ell_\mu(\mathbf{f}^t|\mathbf{N})}{\partial f_{xy}^t}\right), \quad (6)$$

where  $t$  is the iteration number. The “max” operation keeps  $f_{xy}$  nonnegative. The positive variable  $\eta_t$  controls the step size and a simple line search of  $\eta_t$  accelerated the convergence largely.

The initial value is generally important for gradient based algorithms. But for the SPR method, the influence is effectively reduced by controlling  $\mu$ . The MAP estimate is first computed for a large  $\mu$  and then for smaller  $\mu$ 's successively. For a very large  $\mu$ , most components of the optimal  $\mathbf{f}$  are 0 and the influence of the initial value is negligible. The MAP estimate is then used for the initial value of a slightly smaller  $\mu$ , which is a good starting point in general.

## 2.4 HIO method and Bayesian framework

In this subsection, we discuss the relation between the widely used HIO method and the SPR method.

The HIO method [18–20] exploits the knowledge that the positive density is localized in a small support region. Let us define the support region  $\gamma$ , and the method is written as follows,

$$\text{minimize } \sum_{uv} (|F_{uv}|c_{uv}^{1/2} - N_{uv}^{1/2})^2, \text{ subject to } f_{xy} \geq 0, (x, y) \in \gamma \text{ and } f_{xy} = 0, (x, y) \notin \gamma. \quad (7)$$

This problem can be formulated as the MAP estimation problem by setting the Gaussian likelihood function and the uniform prior distribution, that is,

$$p_G(\mathbf{N}|\mathbf{f}) = \frac{1}{(2\pi\sigma^2)^{M^2/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{uv} (|F_{uv}|c_{uv}^{1/2} - N_{uv}^{1/2})^2\right), \quad \tilde{p}_u(f_{xy}) = \begin{cases} C, & (x, y) \in \gamma \\ 0, & (x, y) \notin \gamma \end{cases},$$

where  $C$  is a positive number. If  $f_{xy}$  is unbounded, above  $\tilde{p}_u(f_{xy})$  is improper. The MAP estimate is computed by solving Eq.(7).

Formally, the difference between the HIO method and the SPR method is the likelihood and the prior. But more importantly, the proposed SPR method does not require the support region. In the SPR method, a lot of components of the estimated electron density  $\mathbf{f}$  become 0 automatically.

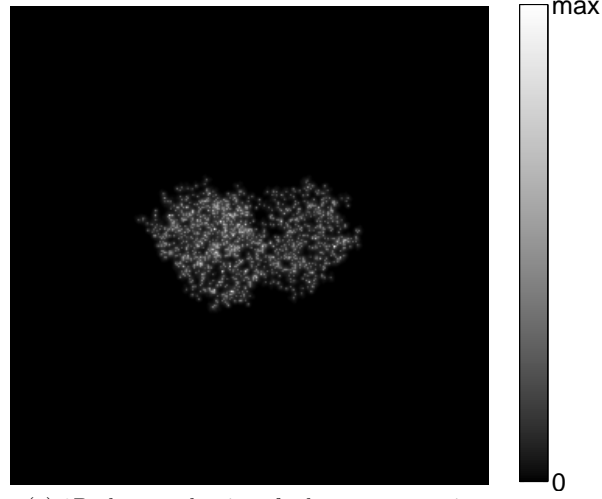
## 3 Numerical Experiment

### 3.1 Simulated data

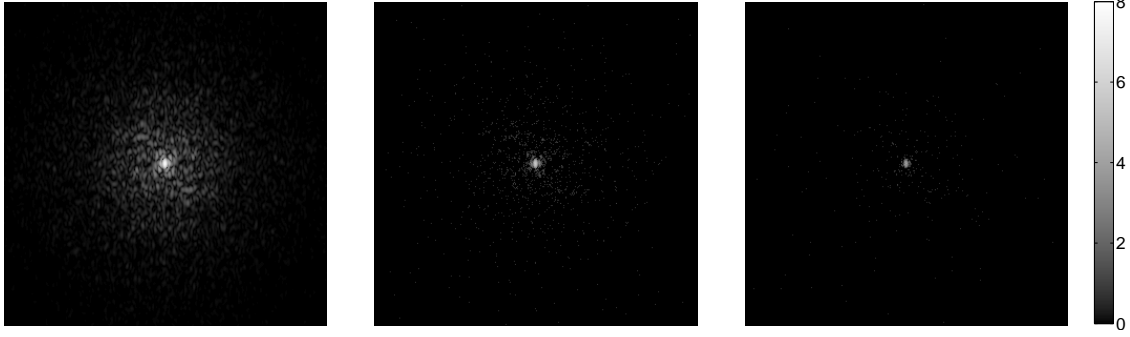
The SPR method was tested with simulated diffraction data. The electron density of lysozyme, a 14 kD protein, was projected onto a 2D plane. The projected density was then Fourier transformed into reciprocal space  $(u, v)$  to simulate a diffraction image. The density and the ideal diffraction image are shown in Figs. 3a and 3b, respectively.

Experimental diffraction images were simulated by converting the diffraction intensity into the number of photons per effective pixel  $(\lambda/\sigma L)^2$  according to the Poisson distribution with the mean,  $S_{uv}$ , defined as follows,

$$S_{uv} = \alpha |F_{uv}|^2 c_{uv} = I r_c^2 \left( \frac{\lambda}{\sigma L} \right)^2 |F_{uv}|^2 c_{uv}.$$



(a) 2D electron density of a lysozyme protein.



(b) A noiseless diffraction image with x-ray fluxes  $I = 5.0 \times 10^{22}$ . The scale of the image is proportional to  $\log(S_{uv} + 1)$ .

(c) Simulated diffraction image with x-ray fluxes  $I = 5.0 \times 10^{21}$ . The total number of photons is 4630. The scale of the image is proportional to  $\log(N_{uv} + 1)$ .

(d) Simulated diffraction image with x-ray fluxes  $I = 1.0 \times 10^{21}$ . The total number of photons is 957. The scale of the image is proportional to  $\log(N_{uv} + 1)$ .

Figure 3: Simulated 2D electron density (a), the diffraction image without noise (b), and simulated diffraction patterns (c,d) of a lysozyme protein are shown. Each image is  $308 \times 308$  pixels. The spatial frequency of the diffraction patterns is  $0.0812 \text{ nm}^{-1}$  and the unit of the x-ray fluxes  $I$  is photons/pulse/ $\text{mm}^2$ .

Here,  $I$  is the incident x-ray flux (photons/pulse/mm<sup>2</sup>),  $r_c$  is the classical electron radius ( $2.82 \times 10^{-12}$  mm),  $\lambda$  is the x-ray wave length (0.1 nm),  $\sigma$  is the oversampling ratio per one dimension (2), and  $L$  is the molecule diameter (6.16 nm) (see Fig. 1). The Ewald sphere curvature was ignored but this is a good approximation for small angles.

Two examples for scattered photons of x-ray fluxes with  $I = 5.0 \times 10^{21}$  and  $1.0 \times 10^{21}$  photons/pulse/mm<sup>2</sup> are shown in Figs. 3c and 3d, respectively. Although the fluxes in the above simulation are around 100 times larger than that of the current LCLS, the sizes of the molecules of interest are generally larger than lysozyme. When the size is around 1000 kD, the total number of photons arriving at the 2D plane will be in the similar range as Figs. 3c and 3d.

### 3.2 Results

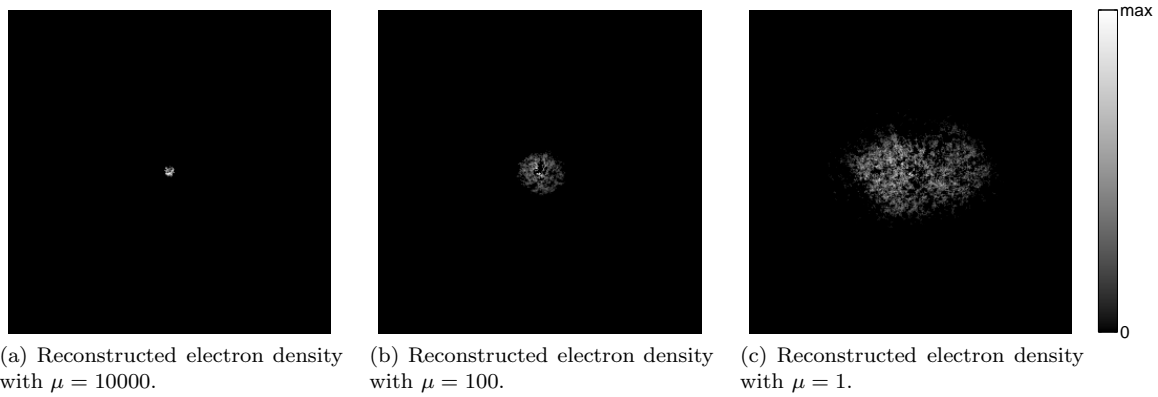


Figure 4: Reconstruction of the electron density from the ideal diffraction image by the SPR method with different values of  $\mu$ .

The SPR method was applied to the data to reconstruct the density. Figure 4 shows the results of the SPR method applied to the ideal diffraction image. As  $\mu$  decreases, more components become positive. The reconstructed electron density becomes similar to the true density when  $\mu = 1$ . Thus, it is important to select the best  $\mu$ . A simple strategy is to rank the results with a criterion. Many criteria have been proposed in statistics, such as, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) [22], but here we use a different one. Let  $\hat{\mathbf{f}}(\mu)$  be the maximizer of  $\ell_\mu(\mathbf{f}|\mathbf{N})$  and the following error is used as the criterion

$$\text{Error}_F(\mu) = \frac{\sum_{uv} (|\hat{F}(\mu)_{uv}| c_{uv}^{1/2} - N_{uv}^{1/2})^2}{\sum_{u'v'} N_{u'v'}}, \quad (8)$$

where  $\hat{F}(\mu)_{uv} = \mathcal{F}(\hat{\mathbf{f}}(\mu))_{uv}$ . This error function has been widely used for the phase retrieval problem and is employed here as well in order to rank the results. Roughly speaking, the error decreases as  $\mu$  shrinks since it controls the balance between the prior and the likelihood. But the error does not necessarily decrease monotonically because the log posterior in Eq. (4) is different from the error. Thus, we choose the  $\mu$  which minimizes  $\text{Error}_F(\mu)$ .

Figure 5 shows the  $\text{Error}_F(\mu)$  for diffraction images in Fig. 3. The value of  $\mu$  was varied from  $1 \times 10^4$  to  $1 \times 10^{-4}$ . Note that although the horizontal axis of Fig. 5 increases from left to right, the value of  $\mu$  was decreased in the experiments. The best  $\mu$  was selected for each diffraction image by  $\text{Error}_F(\mu)$  and the reconstructed electron densities are shown in Figs. 6a, 6b, and 6c.

We also applied the HIO method for the data. The central square ( $154 \times 154$  pixels) was set as the support region. The method has a problem with convergence, and we stopped the iteration loop when the incremental difference of the estimated density becomes small enough. Figure 7 shows the reconstructions

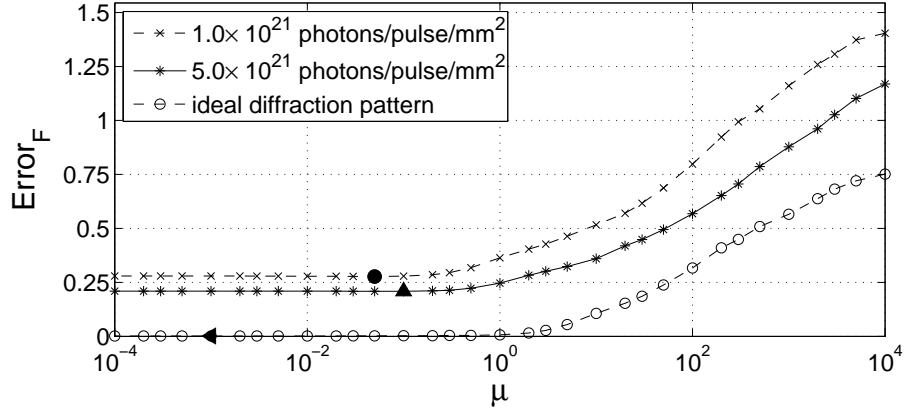


Figure 5: The values of  $\text{Error}_F(\mu)$  for simulated diffraction images in Fig. 3. The black circle and triangles on the curves correspond to the reconstructed density images of Figs. 6.

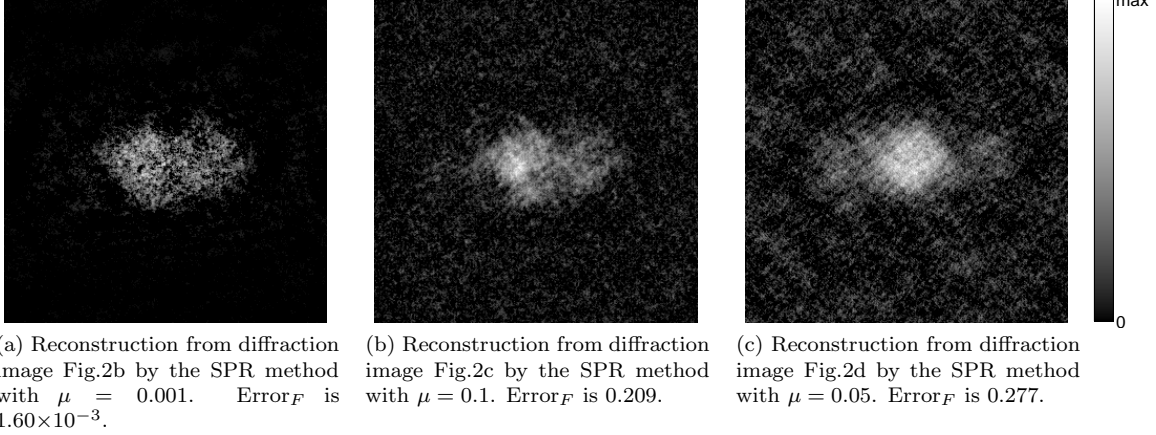


Figure 6: Reconstruction of the 2D electron density by the SPR method.

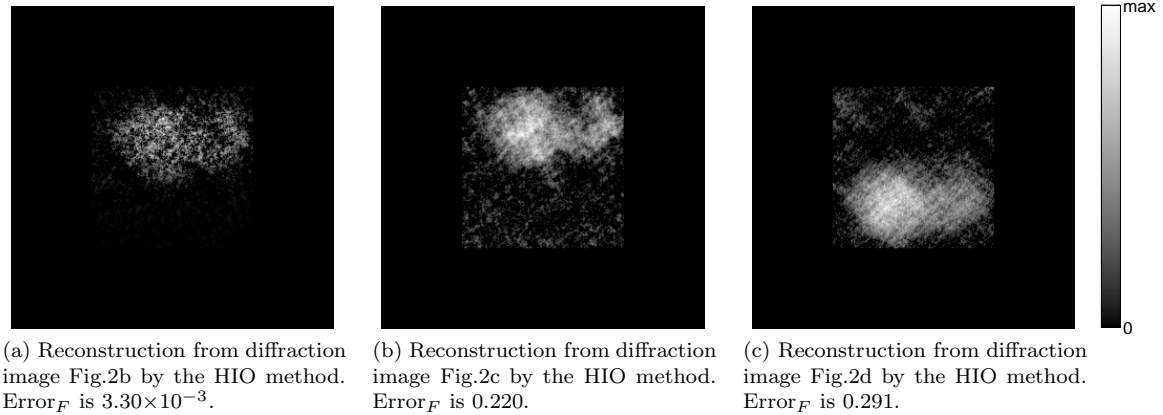


Figure 7: Reconstruction of the 2D electron density by the HIO method.

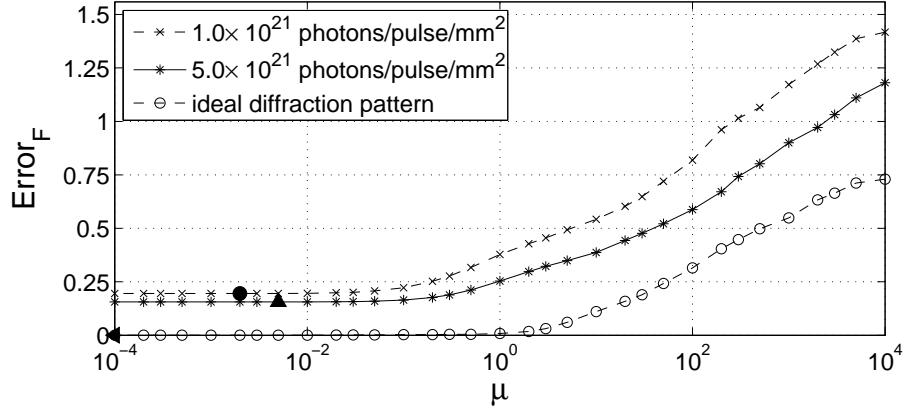


Figure 8: The values of  $\text{Error}_F(\mu)$  for simulated diffraction images in Fig. 3 without the central pixel. The black circle and triangles on the curves correspond to the reconstructed density images of Fig. 9.

obtained by the HIO method for the diffraction images in Fig. 3. The reconstructed electron densities are similar to Fig. 6, but the SPR method obtained better  $\text{Error}_F(\mu)$  for all of the diffraction images.

We note the computational time of the SPR method. The method was implemented with C and run on a desktop machine with an Intel Xeon X5570 (2.93 GHz 4 cores). The optimal electron density was computed for all the values of  $\mu$ 's (33 different values of  $\mu$ ) within 10 minutes.

### 3.3 Missing centers

One of the problems of the phase retrieval from diffraction images is missing centers. The intensities around the center cannot be measured experimentally because the scattered beam is superimposed by the direct beam around the center and a beam stop is used to block the beam.

The SPR method can be applied to the case easily by modifying eq. (4) as follows,

$$\ell_\mu(\mathbf{f}|\mathbf{N}) = \sum_{uv \in \alpha} (N_{uv} \ln |F_{uv}|^2 - |F_{uv}|^2 c_{uv}) - \mu \sum_{xy} w_{xy} f_{xy}, \quad (9)$$

where  $\alpha$  is the active pixel set where the central pixels are excluded. The error function is also modified as

$$\text{Error}_F(\mu) = \frac{\sum_{uv \in \alpha} (|\hat{F}(\mu)_{uv}| c_{uv}^{1/2} - N_{uv}^{1/2})^2}{\sum_{u'v' \in \alpha} N_{u'v'}}. \quad (10)$$

Firstly, we masked the central pixel from each diffraction image in Fig. 3 and applied the SPR method. By masking the central pixel, Fig. 3b lost around 5.7% of the photons, Fig. 3c lost 294 out of 4630 photons and Fig. 3d lost 56 out of 957 photons. Figure 8 shows the errors. Note that the error values are better than those in Fig. 5, but it does not mean the reconstructed densities in Fig. 9 are better than those in Fig. 6, because the errors are computed only for the active pixel set. Figure 10 shows the results of the HIO method. Since the central pixel corresponds to a scaling factor of the electron density of the protein, the reconstructed images are quite similar to those reconstructed from the complete diffraction images. We see that the error values of the SPR method are better than those of the HIO method.

Next, we further masked the central  $3 \times 3$  pixels from each diffraction image and applied the SPR method. Note that Fig. 3b lost around 35% of the photons, Fig. 3c lost 1595 out of 4630 photons and Fig. 3d lost 312 out of 957 photons. More than 30% of the photons are lost and the results are expected to be degraded largely. Figure 11 shows the errors. The reconstructed densities are shown in Fig. 12. Figure 13 shows the results of HIO method. We see that the error values of the SPR method are better and the reconstructed results are more stable than the HIO results. It is worth noting that the SPR

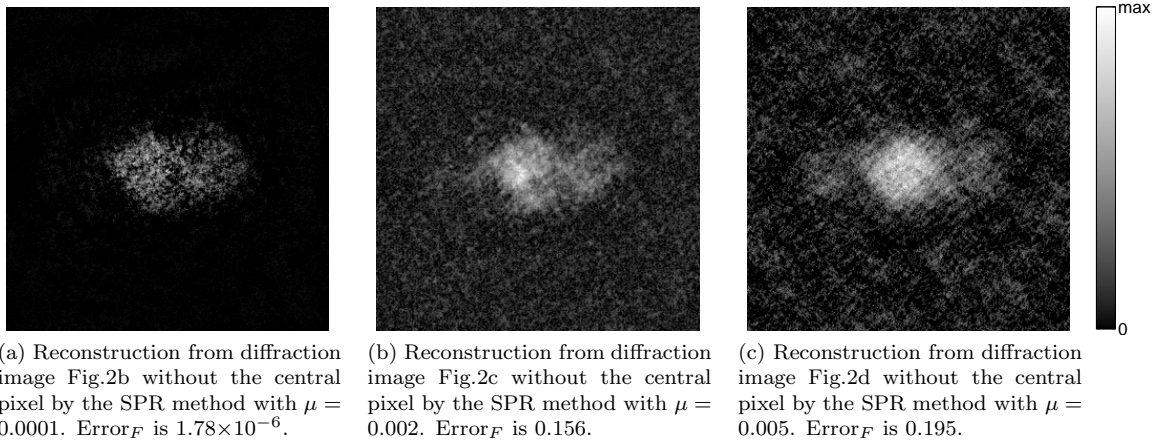


Figure 9: Reconstruction of the 2D electron density by the SPR method.

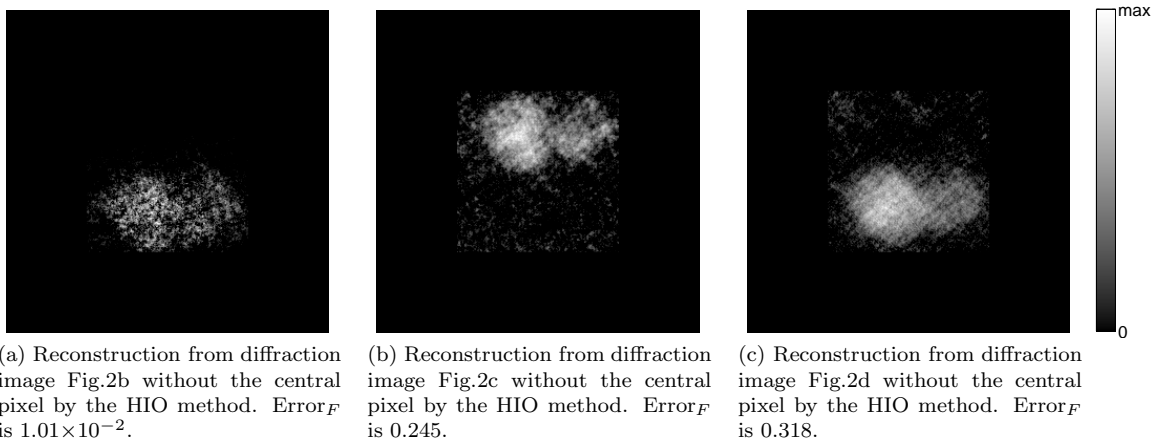


Figure 10: Reconstruction of the 2D electron density by the HIO method.

method worked well even if the central  $5 \times 5$  pixels, a half area of the centrospeckle pattern, were masked. In that condition, 54.4 % of the total photons are lost.

## 4 Discussion

We have proposed a new phase retrieval method. The SPR method is based on the MAP estimation of the Bayesian statistics, it will be effective for single molecule diffraction images which will be obtained by XFELs.

From a Bayesian viewpoint, the difference between other phase retrieval methods including the HIO method and the SPR method is summarized in the likelihood and the prior. Many methods use the squared loss, which corresponds to a Gaussian likelihood, while the present method employs a Poisson distribution. This is suitable for CXDI data because the number of the scattered photons detected at each pixel is a counting process. Similar idea has been found in a related work [16]. The prior in HIO method corresponds to a uniform distribution with a bounded support, while the proposed method uses the exponential distribution. Similar prior has been widely used in statistics [13]. This combination shows a new promising direction for phase retrieval.

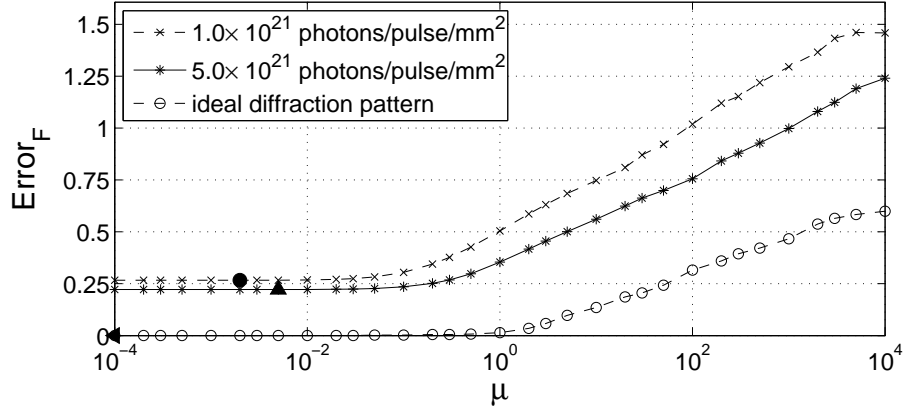


Figure 11: The values of  $\text{Error}_F(\mu)$  for simulated diffraction images in Fig. 3 without the central  $3 \times 3$  pixels. The black circle and triangles on the curves correspond to the reconstructed density images of Figs. 12.

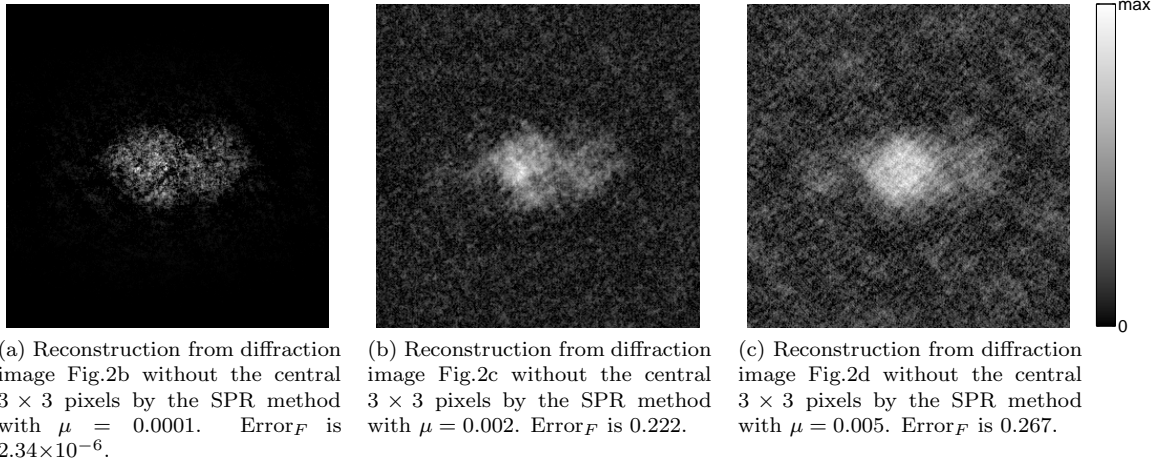


Figure 12: Reconstruction of the 2D electron density by the SPR method.

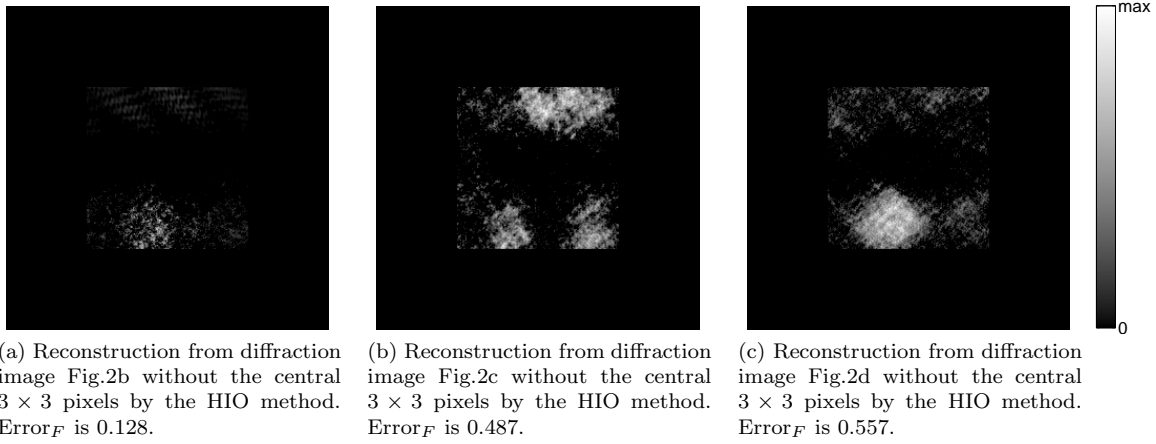


Figure 13: Reconstruction of the 2D electron density by the HIO method.

The proposed method has been tested with simulated data. The electron densities are reconstructed with a reasonable computational cost, and errors of the results are smaller than the HIO method. We further tested the method with the diffraction data where some central pixels are masked. The proposed method is applied without any major modification, and the results are quite stable even some pixels are masked.

The SPR method reconstructs the density from an incomplete observation utilizing the sparsity. This idea is similar to CS [21]. In [15], CS is applied to the phase retrieval problem, however, in our problem, the number counts are very limited, and we needed to extend the model.

Finally, we list some of our future works to improve the SPR method. One is the hyperparameter. The hyperparameter  $\rho(\mu)_{xy}$  can be modified to reflect the knowledge of the true density. This may result in a better estimate. Another issue is the algorithm. The densities for the whole sequence of  $\mu$  in Fig. 5 are computed with a simple line search algorithm but advanced optimization methods may speed it up.

## Acknowledgments

We would like to thank Nobuhiro Go for fruitful discussion and Atsushi Tokuhisa for providing the electron density data. We are also grateful to Satoshi Ito for helpful discussion. This work was partly supported by X-ray Free Electron Laser Utilization Research Project of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

- [1] D. Sayre. Prospects for long-wavelength X-ray microscopy and diffraction. In M. Schlenker, M. Fink, J. P. Goedgebuer, C. Malgrange, J. Ch. Viénot, and R. H. Wade, editors, *Imaging Processes and Coherence in Physics*, volume 112 of *Springer Lecture Notes in Physics*, pages 229–235. Springer, Berlin, 1980.
- [2] K. J. Gaffney and H. N. Chapman. Imaging atomic structure and dynamics with ultrafast x-ray scattering. *Science*, 316:1444–1448, 2007.
- [3] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu. Potential for biomolecular imaging with femtosecond X-ray pulses). *Nature*, 406:752–757, August 2000.
- [4] G. Huldtt, A. Szóke, and J. Hajdu. Diffraction imaging of single particles and biomolecules. *J. Str. Biol.*, 144:219–227, 2003.
- [5] M. Marvin Seibert *et al.* Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature*, 470:78–81, February 2011.
- [6] G. Bortel and G. Faigel. Classification of continuous diffraction patterns: A numerical study. *J. Str. Biol.*, 158(1):10–18, April 2007.
- [7] G. Bortel, G. Faigel, and M. Tegze. Classification and averaging of random orientation single macromolecular diffraction patterns at atomic resolution. *J. Str. Biol.*, 166(3):226–233, 2009.
- [8] R. Fung, V. Shneerson, D. K. Saldin, and A. Ourmazd. Structure from fleeting illumination of faint spinning objects in flight. *Nature Physics*, 5(1):64–67, January 2009.
- [9] N.-T. D. Loh and V. Elser. Reconstructon algorithm for single-particle diffraction imaging experiments. *Physical Review E*, 80:026705, 2009.
- [10] D. K. Saldin, V. L. Shneerson, R. Fung, and A. Ourmazd. Structure of isolated biomolecules obtained from ultrashort x-ray pulses: exploiting the symmetry of random orientations. *J. of Physics: Condensed Matter*, 21(13):134014, 2009.

- [11] J. Frank. *Three-Dimensional Electron Microscopy Of Macromolecular Assemblies: Visualization Of Biological Molecules In Their Native State*. Oxford University Press, Oxford, 2006.
- [12] L. Young *et al.* Femtosecond electronic response of atoms to ultra-intense X-rays. *Nature*, 466:56–61, 2010.
- [13] R. Tibshirani. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc., Ser. B*, 58(1):267–288, 1996.
- [14] P. Zhao and B. Yu. On model selection consistency of lasso. *J. Machine Learning Research*, 7:2541–2563, 2006.
- [15] S. Marchesini. Ab initio compressive phase retrieval. *ArXiv Physics Optics e-prints*, 2008.
- [16] R. Irwan and R. G. Lane. Phase retrieval with prior information. *J. Opt. Soc. Am. A*, 15(9):2302–2311, 1998.
- [17] S. Baskaran and R. P. Millane. Bayesian image reconstruction from partial image and aliased spectral intensity data. *IEEE trans. Image Proc.*, 8(10):1420–1434, 1999.
- [18] J. R. Fienup. Reconstruction of an object from the modulus of its Fourier transform. *Optics Letters*, 3(1):27–29, 1978.
- [19] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21(15):2758–2769, August 1982.
- [20] J. Miao, D. Sayre, and H. N. Chapman. Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects. *J. Opt. Soc. Am. A*, 15(6):1662–1669, 1998.
- [21] D. Donoho. Compressed sensing. *IEEE transaction on Information Theory*, 52(4):1289–1306, April 2006.
- [22] H. Zou, T. Hastie, and R. Tibshirani. On the “degrees of freedom” of the lasso. *Annals of Statistics*, 35(5):2173–2192, 2007.